Méthodologie

Validité d'un test diagnostique : utilité clinique de la sensibilité, spécificité et rapports de vraisemblance

Validity of diagnostic tests : utility of sensitivity, specificity and likelihood ratios values in clinical practice

GUILLAUME CHRISTE, PT, MSc, DAS EN THÉRAPIE MANUELLE

Haute Ecole de Sante Vaud, Filière Physiothérapie, Lausanne

L'auteur atteste ne pas avoir de confits d'intérêts dans la réalisation de ce travail

Keywords

Validity, diagnostic test, clinical reasoning, sensitivity, specificity, likelihood ratio

Mots clés

Validité, test diagnostique, raisonnement clinique, sensibilité, spécificité, rapport de vraisemblance

Abstract

For health professionals who mainly use clinical data to guide decision making process, the diagnostic process may be made up of uncertainty. Diagnostic tests have the aim of reducing the clinical uncertainty by their ability to identify or exclude a particular diagnosis. Knowing the validity of diagnostic tests, with the sensitivity, specificity and likelihood ratios, is an essential step in the clinical reasoning process. Without knowledge of the validity, one cannot interpret adequately the test result and its influence in the exclusion or inclusion of one particular diagnosis. The aim of this article is to define the sensitivity, specificity and likelihood ratios of diagnostic tests. A particular focus of this article is on how the knowledge of these values could influence the clinical reasoning process.

Résumé

Pour des professions utilisant essentiellement des données cliniques afin de guider la prise de décision, comme l'anamnèse et l'examen physique, le processus diagnostique est un chemin qui peut être semé d'embuches et d'incertitudes. Les tests diagnostiques ont pour but de diminuer l'incertitude du diagnostic par leur capacité à identifier ou exclure un diagnostic particulier. Connaître la validité d'un test diagnostique, avec ses valeurs de sensibilité, de spécificité et les rapports de vraisemblance, est une étape indispensable dans le processus de raisonnement clinique. Sans connaissance de ceux-ci, il sera difficile d'interpréter adéquatement le résultat du test et de déterminer son influence dans la décision d'inclure ou d'exclure un diagnostic particulier. Cet article a pour but de définir les notions de sensibilité, spécificité et rapports de vraisemblance d'un test diagnostique. L'accent est mis sur l'utilisation de ces valeurs dans un contexte clinique afin de guider le processus de raisonnement clinique.

Introduction

Chaque clinicien prend quotidiennement des décisions basées sur de nombreuses informations afin de poser un diagnostic (médical, physiothérapeutique ou ostéopathique, entre autres) [1–3]. Des informations récoltées lors de l'anamnèse permettent tout d'abord d'élaborer une série d'hypothèses sur le problème du patient. Ces hypothèses seront ensuite

confrontées aux résultats des tests effectués dans l'examen physique [2–4].

Certains tests devraient permettre d'influencer le processus diagnostique par leur capacité à identifier la présence d'une condition particulière [5]. Aucun test ne permettant d'émettre un diagnostic avec 100% de certitude, le clinicien est confronté à l'usage des probabilités [6]. Un « bon » test sera un test qui augmente beaucoup les probabilités d'avoir ou pas un certain diagnostic, alors qu'un test inefficace n'influencera pas du tout les probabilités d'avoir un diagnostic particulier. C'est la notion de validité d'un test qui définit sa capacité à déterminer la présence ou l'absence d'une condition particulière. Elle se caractérise par ses valeurs de sensibilité, de spécificité et des rapports de vraisemblance *(likelihood ratio)* [7,8]. Cet article a pour objectif de définir ces notions et de donner des exemples afin de les décrire dans une perspective clinique. La validité se différencie de la fiabilité des tests (intra et inter observateur, par exemple) qui ne sera pas abordée ici.

Développement

La sensibilité et la spécificité d'un test permettent de connaître l'exactitude *(« accuracy »)* diagnostique d'un test. Plus ces valeurs se rapprochent de 100%, plus le test est sensible et spécifique. Le *likelihood ratio* est une valeur complémentaire à celle de la sensibilité et spécificité qui permet de répondre à la question de savoir qu'elles sont les chances qu'un individu ait ou n'ait pas la maladie en fonction du résultat du test [8,9]. La sensibilité, la spécificité et le *likelihood ratio* sont des valeurs intrinsèques à chaque test diagnostique et ne prennent pas en compte la prévalence de la maladie dans la population [8]. D'autres notions, comme les valeurs prédictives positives et négatives prennent en compte la prévalence, mais ne seront pas abordées ici [10].

Sensibilité

La sensibilité d'un test permet de déterminer sa capacité à détecter les vrais positifs, c'est-à-dire les personnes qui sont atteintes de la maladie. Elle se calcule dans un échantillon de personnes chez qui le gold standard est positif [11]. Le gold standard est le test de référence pour détecter une maladie. La proportion des patients dont le test est positif donne la valeur de sensibilité. Par exemple, chez une population de 100 personnes avec une rupture complète du ligament croisé antérieur (LCA) confirmée par arthroscopie ou par imagerie par résonance magnétique (IRM), le test de *Lachman* est positif chez 96 patients, alors qu'il est négatif chez 4 patients [12]. En prenant le calcul présenté sur la figure 1A, cela donne une sensibilité de 96%, ce qui est excellent.

Un test avec une bonne sensibilité permet de détecter facilement les porteurs de la maladie et est particulièrement utile pour **exclure une pathologie dans le cas où il est négatif *(« rule out »)***. Par contre, comme il a beaucoup de chance de détecter des faux positifs, un résultat positif n'aura pas beaucoup d'utilité dans le processus diagnostique. En prenant l'exemple ci-dessus, on peut déduire qu'un test de *Lachman* négatif est efficace pour exclure une déchirure totale du LCA.

Spécificité

La spécificité d'un test détermine sa capacité à détecter les vrais négatifs, c'est-à-dire les personnes non atteintes de la maladie. Celle-ci s'apprécie chez une population de personnes non atteintes de la maladie, définie par le test gold standard. Reprenons l'exemple du test de *Lachman*. Sur une population de 100 personnes sans lésion du LCA, le test de *Lachman* est négatif 78 fois. Cela veut dire que sur les 100 individus testés, 22 ont un test faussement positif et 78 un test négatif à juste titre. En utilisant le calcul de la figure 1B, il est possible de déterminer la spécificité du test de *Lachman* à 78% [12].

Un test avec une bonne spécificité est un test qui permet de détecter correctement les non porteurs de la maladie, il est donc particulièrement efficace pour **confirmer la présence d'une pathologie lorsqu'il est positif *(« rule in »)***. Par contre, comme il peut rater de nombreux vrais positifs, ce n'est pas un bon test d'exclusion lorsqu'il est négatif. Au vu de sa meilleure sensibilité que spécificité, le test de *Lachman* a plus de chances d'être correct lorsqu'il est négatif que quand il est positif. C'est donc un meilleur test pour exclure que pour confirmer une rupture du LCA.

Likelihood ratios

Les *likelihood ratios* (rapport de vraisemblance) permettent de répondre à la question : « Dans quelle mesure la probabilité que ce patient ait cette pathologie est augmentée (ou diminuée) si son test est positif (ou négatif) ? » [8]. Un *likelihood ratio* positif (LR+), plus haut que 1, indique le coefficient par lequel la probabilité d'avoir la pathologie augmente si le résultat du test est positif (par rapport à la probabilité pré-test) [13]. Un LR+ supérieur à 5 est considéré comme des chances modérées et un LR+ supérieur à 10 comme de fortes chances que la pathologie soit présente. Le *likelihood ratio* négatif (LR-) représente le coefficient par lequel la probabilité d'avoir la pathologie diminue lorsque le test est négatif (par rapport à la probabilité pré-test). Le LR- est compris entre 0 et 1 et plus il est petit, plus les chances de ne pas avoir la maladie augmentent [13]. Un LR- entre 0.1 et 0.2 indique des chances modérées et un LR- inférieur à 0.1 de fortes chances que la pathologie soit absente. Les *likelihood ratios* sont calculés sur la base des valeurs de sensibilité et spécificité (Figure 1C).

Une revue systématique récente sur la validité des tests méniscaux du genou suggère que le test de *McMurray's* a un LR+ de 3.2 et un LR- de 0.52 [13]. Cela signifie que lorsque le test est positif, le patient a environ trois fois plus de chance d'avoir une déchirure méniscale, ce qui reste une petite probabilité d'avoir la pathologie. La valeur de LR- montre que si le test est négatif, les chances de ne pas avoir la pathologie ne sont pas plus fortes, ce qui revient à dire que ce test n'a aucune utilité pour exclure une déchirure méniscale. Cette revue confirme la faible validité du test de *McMurray's* [14–17], particulièrement pour exclure une déchirure méniscale lorsqu'il est négatif. Sur la base des *likelihood ratios*, un résultat négatif à ce test ne va devrait pas être considéré dans le processus diagnostique.

A $$\text{Sensibilité} = \frac{\text{Nombre de vrais positifs}}{\text{Nombre de vrais positifs} + \text{Nombre de faux négatifs}}$$

B $$\text{Spécificité} = \frac{\text{Nombre de vrais négatifs}}{\text{Nombre de vrais négatifs} + \text{Nombre de faux positifs}}$$

C $$\text{LR+} = \frac{\text{Sensibilité}}{\text{1 - spécificité}} \qquad \text{LR-} = \frac{\text{1 - Sensibilité}}{\text{Spécificité}}$$

› Figure 1 : calcul de la sensibilité (A), spécificité (B), *likelihood ratio* positif (LR+) et négatif (LR-)

Exemples

Diagnostic d'exclusion d'une myélopathie cervicale

Il est maintenant reconnu que la majorité des cervicalgies peuvent être considérées comme des affections médicales bénignes ou non-spécifiques. Il est toutefois recommandé par la majorité des guidelines de réaliser un *screening* de pathologies sérieuses lors de l'évaluation [18]. Les tests utilisés lors de ce processus de *screening* devraient idéalement avoir une sensibilité élevée et un LR- bas afin d'avoir le plus de certitude que la pathologie n'est pas présente quand le test est négatif.

Parmi les pathologies sérieuses, on trouve une atteinte de la moelle épinière, nommée myélopathie [18–20]. Le Babinski, l'hyperréflexie et le clonus ont été souvent décrits comme des tests de *screening* permettant d'exclure une telle affection [21]. Les valeurs de sensibilité se situent pour le Babinski entre 7 et 53%, pour l'hyperréflexie entre 15 et 44% (biceps, triceps et tendon achilléen) et pour le clonus entre 7 et 35%, suivant les études. Leurs LR- se situent entre 0.7 et 0.94 [20]. En d'autres termes, lorsqu'ils sont négatifs, la probabilité de ne pas avoir une myélopathie cervicale n'est pas diminuée et les chances de ne pas détecter de vrais positifs sont importantes. De l'autre côté, les valeurs de spécificité du Babinski sont entre 92 et 100%, de l'hyperréflexie entre 71 et 98% et du clonus entre 96 et 100% [20]. Contrairement à ce que l'on pense couramment, ce sont de bons tests d'inclusion *(« rule in »)* pour permettre de confirmer la présence d'une myélopathie et des mauvais tests d'exclusion *(« rule out »)*.

Une possibilité pour améliorer le diagnostic est de combiner les tests et les informations subjectives. Par exemple, *Cook et al.* (2011) ont montré qu'une combinaison de 5 critères améliorait fortement les chances d'inclure ou d'exclure une myélopathie cervicale [20]. Ces critères sont : une instabilité à la marche, un test de *Hoffmann* positif, un signe du supinateur inversé *(inverted supinator sign)*, un test de *Babinski* positif et avoir plus de 45 ans. Lorsqu'aucune ou une seule de ces informations est présente, la sensibilité est de 94% et le LR- est de 0.18. Il y a donc des chances modérées que le patient ne soit pas atteint d'une myélopathie et cela peut être une aide importante au diagnostic. Dans le cas où 3 tests (ou plus) sont positifs, la spécificité est de 99% et le LR+ de 30.9. Ceci signifie qu'il y a 30 fois plus de chances que le patient ait une myélopathie.

En conclusion, la connaissance de la validité des tests traditionnellement utilisés dans le processus de *screening* d'une myélopathie est essentielle. En effet, si ces tests sont utilisés de manière isolée, une réponse négative ne permet pas d'exclure une myélopathie. Dès lors, si d'autres informations suggèrent une pathologie spécifique du rachis, le patient sera quand même référé chez le médecin malgré la réponse négative à ces tests [22]. Par contre, la plupart de ces tests sont très utiles pour confirmer la présence d'une myélopathie lorsqu'ils sont positifs, au vu de leur excellente spécificité. Finalement, la combinaison de différentes informations subjectives et objectives peut être d'une grande aide au diagnostic.

Conflit sous-acromial et tendinopathie de la coiffe des rotateurs

Lors de la démarche diagnostique visant à identifier un conflit sous-acromial ou une atteinte de la coiffe des rotateurs, des tests diagnostiques sont fréquemment utilisés. La validité de ces tests a été étudiée à de nombreuses reprises et plusieurs revues ont été publiées à ce sujet [23–25].

Il en ressort que la plupart des tests ont une validité si médiocre qu'ils ont peu d'utilité pour aider au diagnostic [26,27]. Cependant, en analysant leurs valeurs de sensibilité et spécificité, il est quand même possible de proposer certaines recommandations pour leur utilisation clinique. Premièrement, la spécificité de la plupart des tests destinés à identifier un conflit sous-acromial ou une tendinopathie de la coiffe des rotateurs est basse, ce qui rend leur utilisation particulièrement difficile pour identifier *(« rule in »)* une structure particulière (par exemple un tendon spécifique) ou un diagnostic particulier. Par exemple, les tests de *Hawkins Kennedy* ou de *Neer*, avec une spécificité de 57% et 58% et un LR+ de 1.7 et 1.86, respectivement, ne permettent pas d'augmenter les probabilités d'un conflit sous-acromial lorsqu'ils sont positifs [23,28]. Par contre, leur sensibilité est plus élevée, rendant intéressant les tests d'exclusion d'un conflit sous-acromial *(« rule out »)* lorsqu'ils sont né-

Nom du test	Procédure
Genou	
Lachman	Patient en DD, avec le genou fléchi entre 10 et 20°. Une main stabilise le fémur et l'autre main effectue une traction antérieure du tibia. Le test est positif lors de la présence d'un tiroir antérieur et d'un arrêt souple [32].
McMurray's	En partant d'une position de flexion de genou, l'examinateur applique une force de varus et une rotation médiale pendant l'extension du genou. La manœuvre est répétée en valgus et rotation latérale. Le test est positif si il y a une douleur ou un claquement [33].
Myélopathie cervicale	
Babinski	Patient en DD, le thérapeute applique une stimulation sur la partie plantaire du pied (de latéral à médial et du talon aux orteils). Le test est positif si l'hallux part en extension et les orteils en abduction [34].
Clonus	Patient assis ou en DD, le thérapeute fait un étirement rapide du triceps sural en faisant une flexion dorsale rapide de la cheville. Une réponse positive à ce test est une réaction avec plusieurs contractions (saccades) du triceps sural [34].
Hoffmann	Patient assis, le thérapeute stabilise l'articulation inter-phalangienne proximale du 3e doigt et pince rapidement l'ongle vers la flexion entre son pouce et son index. Le test est positif si le pouce du patient fait une abduction et si les doigts une flexion [34].
Supinateur inversé	Patient assis, avec l'avant-bras en pronation et relâché. Avec un marteau réflexe, le thérapeute frappe plusieurs fois l'insertion distale du tendon du brachio-radial proche du processus styloïde du radius. Le test est positif lors d'une flexion des doigts ou une extension du coude [34].
Epaule	
Drop arm	Lâchage lors du retour lent et actif d'une position d'abduction [35].
Hawkins Kennedy	Avec l'épaule à 90° de flexion, une rotation médiale forcée est appliquée en fixant la scapula. Le test est positif s'il reproduit des douleurs [36].
Infra-épineux	Coude fléchi à 90°, bras le long du corps, l'examinateur applique une résistance à la rotation latérale. Le test est positif si faiblesse ou lâchage [29].
Neer	Flexion passive du bras en rotation médiale d'épaule, avec fixation de la scapula. Le test est positif s'il reproduit des douleurs [36].

› Tableau 1 : procédure des tests diagnostiques. DD : décubitus dorsal

Myélopathie cervicale	LR+	LR-
Instabilité à la marche, test de Hoffmann, test de Babinski, test du clonus et signe du supinateur inversé		
Si 3 ou plus présents	30.9	
Si aucun présent		0.18
Rupture de la coiffe des rotateurs		
Arc douloureux, plus de 60 ans, test du drop arm et test de l'infra-épineux (présence ou absence des 4)	28	0.09
Conflit sous acromial		
Hawkins Kennedy, arc douloureux et test de l'infra-épineux (présence ou absence des 3)	10.56	0.17

› Tableau 2 : combinaison de tests pouvant aider au diagnostic. LR+: *Likelihood ratio* positif; LR-: *Likelihood ratio* négatif

gatifs, avec des valeurs de sensibilité pour le *Hawkins Kennedy* et le *Neer* de 74-80% et 72-78%, respectivement [24,28]. Ces deux tests ont donc la possibilité d'influencer le diagnostic uniquement lorsqu'ils sont négatifs, mais pas lorsqu'ils sont positifs.

La combinaison de différents tests peut aussi aider à la prise de décision. Par exemple, quatre critères peuvent être utiles afin d'identifier un patient avec une rupture de la coiffe des rotateurs. Ces critères sont un arc douloureux, avoir plus de 60 ans, un test du drop arm positif et un test positif de l'infra-épineux [25,29]. Si ces informations sont toutes présentes, le LR+ est de 28, ce qui augmente de 28 fois les chances d'avoir une rupture de la coiffe des rotateurs. A l'opposé, si tous ces tests sont négatifs, le LR- est de 0.09, il y a donc une forte probabilité que le patient n'ait pas de rupture de la coiffe des rotateurs. Concernant la présence d'un conflit sous-acromial, la combinaison d'un *Hawkins Kennedy*, d'un arc douloureux et d'un test de l'infra-épineux donne un LR+ de 10.56 quand ils sont tous positifs et un LR- de 0.17 quand ils sont négatifs [25,29]. Cette combinaison est donc aussi utile pour le diagnostic.

Discussion

Les divers exemples utilisés dans cet article mettent en évidence l'importance de connaître les valeurs de sensibilité, spécificité et des *likelihood ratios*. En clinique, un nombre important de tests utilisés ont peu de valeur diagnostique. Certains sont meilleurs pour un processus d'exclusion (lorsqu'ils sont négatifs), d'autres pour confirmer une pathologie (lorsqu'ils sont positifs). La combinaison de certains tests permet parfois d'augmenter le degré de confiance dans le diagnostic, encore faut-il savoir lesquels. Néanmoins, même les meilleurs tests ne sont pas parfaits et doivent être interprétés en regard des autres informations disponibles (anamnèse, mouvements actifs, etc.) afin d'augmenter la probabilité d'être sur le bon diagnostic. Un certain degré d'incertitude fait partie de la pratique clinique [6] et plus les informations récoltées sont de bonne qualité et interprétées adéquatement, plus elles peuvent aider à soutenir un diagnostic.

Les tests de l'épaule ont des valeurs de sensibilité et spécificité qui varient beaucoup suivant les études. C'est pour cela que les valeurs présentées dans cet article sont majoritairement basées sur des revues systématiques ou méta-analyses. Néanmoins, cela soulève quelques questions sur les facteurs pouvant influencer les valeurs de sensibilité et de spécificité. Premièrement, la manière dont les tests sont réalisés, et notamment les critères qui sont utilisés pour déterminer si les tests sont positifs, peut avoir une influence sur leur validité. Par exemple, il peut y avoir des différences importantes dans le calcul de la sensibilité et de la spécificité du test de l'infra-épineux si le critère est la douleur versus le manque de force lors de la rotation latérale contre résistance [23–25]. Un autre facteur qui influence les valeurs de sensibilité et spécificité est la comparaison avec le gold standard de référence [5]. Dans de nombreuses situations, le gold standard n'est pas un test parfait. Ainsi, les tests cliniques pour identifier une tendinopathie de la coiffe des rotateurs sont fréquemment comparés à une vue des tendons lors d'une arthroscopie ou une IRM. Cependant, il est maintenant reconnu que la dégénérescence des tendons de la coiffe est un phénomène naturel apparaissant avec l'âge aussi chez les sujets asymptomatiques. Une personne sur deux avec un conflit sous-acromial présente des signes de pathologie du supra-épineux, ce qui est la même proportion que celle que l'on retrouve chez une population asymptomatique du même âge [26,30]. Finalement, le contexte dans lequel l'étude et les tests sont réalisés peut influencer les résultats. Si l'on sélectionne uniquement dans une étude des patients qui ont une présentation claire de la pathologie et que l'on teste la sensibilité et la spécificité de certains tests, il se peut que les valeurs soient faussement élevées, car tous les patients ayant une présentation plus atypique n'ont pas été inclus [31].

Conclusion

Connaître la validité d'un test diagnostique, avec ses valeurs de sensibilité, de spécificité et les *likelihood ratios*, est une étape indispensable dans le processus de raisonnement clinique. Ainsi, le clinicien pourra apprécier la valeur des informations récoltées et leur influence dans l'établissement du diagnostic. Un test sensible, avec un LR- bas, sera utile lorsqu'il est négatif, afin d'exclure un diagnostic particulier. Un test spécifique, avec un LR+ élevé, sera utilisé afin de confirmer un diagnostic. Utilisés adéquatement, les tests cliniques devraient permettre de diminuer l'incertitude clinique et d'augmenter les probabilités d'être sur le bon diagnostic.

Implications pour la pratique

- L'influence du résultat d'un test sur le processus diagnostique dépend de sa validité.
- Un « bon » test diagnostique permet d'influencer les probabilités d'être en présence d'un certain diagnostic.
- Un test sensible est utile comme test d'exclusion lorsqu'il est négatif.
- Un test spécifique est utile pour confirmer un diagnostic lorsqu'il est positif.

Contact

Guillaume Christe
Guillaume.christe@hesav.ch

Références

1. Christe G. Décision kinésithérapique: épicondylalgie latérale. Kinésithérapie, la Revue. 2015; 15(159): 19–26.
2. Jones MA. Clinical Reasoning for Manual Therapists. Butterworth Heinemann; 2004.
3. Edwards I, Jones M, Carr J, Braunack-Mayer A, Jensen GM. Clinical reasoning strategies in physical therapy. Phys Ther. 2004; 84(4): 312–35.
4. Higgs J, Jones MA, Loftus S, Christensen N. Clinical Reasoning in the Health Professions. 3rd Edition. Butterworth-Heinemann Ltd; 2008.
5. Knottnerus JA, van Weel C, Muris JW. Evaluation of diagnostic procedures. BMJ. 2002; 324(7335): 477–80.
6. Simpkin AL, Schwartzstein RM. Tolerating Uncertainty — The Next Medical Revolution? N Engl J Med. 2016; 375(18): 1713–5.

7. Portney LG, Watkins MP. Foundations of Clinical Research: Applications to Practice. Prentice Hall Health; 2000.
8. Deeks JJ, Altman DG, Altman D, Bland J, Altman D, Bland J, et al. Diagnostic tests 4: likelihood ratios. BMJ. 2004; 329(7458): 168–9.
9. Smith C. Diagnostic tests (1) – sensitivity and specificity. Phlebology. 2012; 27: 250–1.
10. Smith CJ. Diagnostic tests (2) - positive and negative predictive values. Phlebology. 2012; 27(5): 305–6.
11. Nendaz MR, Perrier a. Sensibilité, spécificité, valeur prédictive positive et valeur prédictive négative d'un test diagnostique. Rev Mal Respir. 2004; 21(556358): 390–3.
12. Leblanc M-C, Kowalczuk M, Andruszkiewicz N, Simunovic N, Farrokhyar F, Turnbull TL, et al. Diagnostic accuracy of physical examination for anterior knee instability: a systematic review. Knee Surgery, Sport Traumatol Arthrosc. 2015; 23(10): 2805–13.
13. Smith BE, Thacker D, Crewesmith A, Hall M. Special tests for assessing meniscal tears within the knee: a systematic review and meta-analysis. Evid Based Med. 2015; 20(3): 88–97.
14. Hegedus EJ, Cook CCC, Hasselblad V, Goode A, McCrory DC, Hasselblad FV, et al. Physical examination tests for assessing a torn meniscus in the knee: a systematic review with meta-analysis. J Orthop Sports Phys Ther. 2007; 37(9): 541–50.
15. Scholten RJ, Devillé WL, Opstelten W, Bijl D, van der Plas CG, Bouter LM. The accuracy of physical diagnostic tests for assessing meniscal lesions of the knee: A meta-analysis. J Fam Pract. 2001; 50(11): 938–44.
16. Solomon DH, Simel DL, Bates DW, Katz JN, Schaffer JL. The rational clinical examination. Does this patient have a torn meniscus or ligament of the knee? Value of the physical examination. JAMA. 2001; 286(13): 1610–20.
17. Meserve BB, Cleland J a, Boucher TR. A meta-analysis examining clinical test utilities for assessing meniscal injury. Clin Rehabil. 2008; 22(2): 143–61.
18. Childs JD, Cleland JA, Elliott JM, Teyhen DS, Wainner RS, Whitman JM, et al. Neck pain: Clinical practice guidelines linked to the International Classification of Functioning, Disability, and Health from the Orthopedic Section of the American Physical Therapy Association. J Orthop Sport Phys Ther. 2008; 38(9): A1–34.
19. Verhagen AP, Downie A, Popal N, Maher C, Koes BW. Red flags presented in current low back pain guidelines: a review. Eur Spine J. 2016; 25(9): 2788-802
20. Cook CE, Wilhelm M, Cook AE, Petrosino C, Isaacs R. Clinical tests for screening and diagnosis of cervical spine myelopathy: A systematic review. J Manipulative Physiol Ther. 2011; 34(8): 539–46.
21. Butler D. The sensitive nervous system. Adelaide: Noigroup Publications; 2000.
22. Christe G, Hall T. The screening process of a patient with low back pain and suspected thoracic myelopathy: a case report. J Man Manip Ther. 2017; In press. DOI: 10.1080/10669817.2017.1282189
23. Hegedus EJ, Goode A, Campbell S, Morin A, Tamaddoni M, Moorman CT, et al. Physical examination tests of the shoulder: a systematic review with meta-analysis of individual tests. Br J Sports Med. 2008; 42(2): 80–92
24. Hegedus EJ, Goode AP, Cook CE, Michener L, Myer CA, Myer DM, et al. Which physical examination tests provide clinicians with the most value when examining the shoulder? Update of a systematic review with meta-analysis of individual tests. Br J Sports Med. 2012; 46(14): 964–78.
25. Hegedus EJ, Cook C, Lewis J, Wright A, Park J-Y. Combining orthopedic special tests to improve diagnosis of shoulder pathology. Phys Ther Sport. 2015;16(2):87–92.
26. Lewis J, McCreesh K, Roy J-S, Ginn K. Rotator Cuff Tendinopathy: Navigating the Diagnosis-Management Conundrum. J Orthop Sports Phys Ther. 2015;1–43.
27. Lewis J. Masterclass Rotator cuff related shoulder pain: Assessment, management and uncertainties. Man Ther. 2016;23:57–68.
28. Alqunaee M, Galvin R, Fahey T. Diagnostic accuracy of clinical tests for subacromial impingement syndrome: a systematic review and meta-analysis. Arch Phys Med Rehabil. 2012; 93(2): 229–36.
29. Park H, Yokota A, Gill H, El Rassi G, McFarland E. Diagnostic Accuracy of Clinical Tests for the Different Degrees of Subacromial Impigment Syndrome. J bone Jt Surg. 2005; 87-A(7): 1446–55.
30. Frost P, Andersen JH, Lundorf E. Is supraspinatus pathology as defined by magnetic resonance imaging associated with clinical sign of shoulder impingement? J Shoulder Elbow Surg. 1999; 8(6): 565–8.
31. Pewsner D, Battaglia M, Minder C, Marx A, Bucher HC, Egger M. Information in practice a note of caution. Epidemiology. 2004; 329(7459): 209–13.
32. Cleland J. Examen clinique de l'appareil locomoteur: Tests, évaluation et niveaux de preuve. Elsevier Masson; 2012.
33. Stratford PW, Binkley J. A review of the McMurray test: definition, interpretation, and clinical usefulness. J Orthop Sport Phys Ther. 1995; 22(3): 116–20.
34. Cook C, Roman M, Stewart KM, Leithe LG, Isaacs R. Reliability and diagnostic accuracy of clinical special tests for myelopathy in patients seen for cervical dysfunction. J Orthop Sports Phys Ther. 2009; 39(3): 172–8.
35. Moen MH, de Vos R-J, Ellenbecker TS, Weir A. Clinical tests in shoulder examination: how to perform them. Br J Sports Med. 2010; 44(5): 370–5.
36. Johansson K, Ivarson S. Intra- and interexaminer reliability of four manual shoulder maneuvers used to identify subacromial pain. Man Ther. 2009; 14(2): 231–9.

Quizz (testez vos connaissances)

Pour chacune des cinq questions suivantes, choisissez la réponse qui vous paraît la plus juste.

1. Si un test a une valeur de sensibilité de 48% et une valeur de spécificité de 95%, le risque est de « louper »

a. un vrai positif

b. un vrai négatif

2. Un test avec une sensibilité haute et une spécificité basse donne plus d'information lorsqu'il est positif ou négatif ?

a. positif

b. négatif

3. Un test avec une spécificité élevée aura plus de chance d'avoir

a. Un *likelihood ratio* négatif bas et sera un bon de test d'exclusion

b. Un *likelihood ratio* positif haut et sera un bon de test d'exclusion

c. Un *likelihood ratio* négatif bas et sera un bon de test d'inclusion

d. Un *likelihood ratio* positif haut et sera un bon de test d'inclusion

4. Un patient de 72 ans se présente avec des douleurs d'épaule. Le test de *Hawkins Kennedy* et de *Neer* sont positifs.

a. Il y a de fortes chances que le patient ait un conflit sous-acromial

b. Il y a de fortes chances que le patient ait une déchirure de la coiffe de rotateurs

c. Ces deux tests seuls n'ont aucune utilité pour aider au diagnostic s'ils sont positifs.

Réponses

1. A; 2. B; 3. D; 4. C